スイッチング回帰モデル推定の最尤法と積率母関数法

原 田 桂 一 郎

目　　次

1　はじめに

経済変数間の関係を回帰式で表現し，時系列データを用いて分析を行う場合，データの標本期間にわたって，回帰係数が一定であることを前提とする。標本期間において回帰係数が変化すること，すなわち構造変化が生じる場合には，この基本的前提が崩れることになる。推定期間において固定係数の前提が維持できない場合の回帰モデルとして，時変係数回帰モデル，ランダム係数回帰モデルそしてスイッチング回帰モデルがある。構造変化の模様をどのように仮定するかによって，三者のうちどのモデルを選択するかが決定される。時変係数回帰モデルは，回帰係数が一定の確率過程にしたがって変化することを仮定し，ランダム係数回帰モデルは，ランダムに観測時点間にわたって変化すると仮定する。

スイッチング回帰モデルは標本期間内で構造変化が起こるが，構造変化時点以外のところでは回帰係数が一定と認められる場合のモデルである。通常，変化時点は標本期間内で一つ（一度の構造変化）と仮定し，その時

点の前後のそれぞれの標本区間について回帰式を定式化する。ここで構造変化時点が既知である場合とそうでない場合とを分離して考えてみる。

標本期間（$t=1, 2, \cdots, t_0, t_0+1, \cdots, n$）において，構造変化時点 t_0 が既知の場合，この時点を境にして従属変数ベクトル $Y(n\times 1)$ と独立変数行列 $X(n\times K)$ は二分され，

$$Y=\begin{bmatrix} Y_1 \\ Y_2 \end{bmatrix}, \quad X=\begin{bmatrix} X_1 \\ X_2 \end{bmatrix},$$

回帰式は次のように書くことができる。

$$\begin{bmatrix} Y_1 \\ Y_2 \end{bmatrix}=\begin{bmatrix} X_1 & 0 \\ 0 & X_2 \end{bmatrix}\begin{bmatrix} \boldsymbol{\beta}_1 \\ \boldsymbol{\beta}_2 \end{bmatrix}+\begin{bmatrix} \boldsymbol{e}_1 \\ \boldsymbol{e}_2 \end{bmatrix} \tag{1-1}$$

ここで，ランダム・ベクトル $\boldsymbol{e}_1$ と $\boldsymbol{e}_2$ はそれぞれ $N(0, \sigma_1{}^2I)$，$N(0, \sigma_2{}^2I)$ にしたがって分布し，また $(\boldsymbol{\beta}_1, \sigma_1{}^2)\neq(\boldsymbol{\beta}_2, \sigma_2{}^2)$ が仮定される。

二つの回帰式はそれぞれ別々に推定されるが，Poirier〔14〕は二つの回帰式を t_0 時点で円滑に結合する方法を提示している。既知である t_0 において両者を結びつけるために，三次の回帰方程式を用いている。その方程式を

$$g_i(t)=a_it^3+b_it^2+c_it+d_i, \quad i=1, 2 \tag{1-2}$$

とすると，回帰モデル（1-1）は次のように書くことができる。

$$y_t=\begin{cases} g_1(t)+e_t, & t=1, \cdots, t_0 \\ g_2(t)+e_t, & t=t_0+1, \cdots, n \end{cases} \tag{1-3}$$

t_0 において二つの回帰式が円滑に結び付けられる条件は，t_0 で回帰方程式が等しくなること，回帰方程式の第一次および第二次導関数が等しくなることである。すなわち，

$$g_1(t_0)=g_2(t_0), \quad g_1{}'(t_0)=g_2{}'(t_0), \quad g_1{}''(t_0)=g_2{}''(t_0).$$

ところで，構造変化の発生時点が不明である場合，その時点 t_0 を推定し，構造変化の統計的有意性を検定することが関心事となる。(1-1) 式において $e_{t1} \sim N(0, \sigma_1^2)$ と $e_{t2} \sim N(0, \sigma_2^2)$ が相互に独立と仮定すると，Goldfeld and Quandt〔7〕は未知数 t_0 の尤度関数を次のように与えている。

$$L(y/t_0) = (2\pi)^{-n/2} \sigma_1^{-t_0} \sigma_2^{-(n-t_0)} \exp\left\{-\frac{1}{2\sigma_1^2} \sum_{t=1}^{t_0} (y_t - \boldsymbol{X}_t' \boldsymbol{\beta}_1) - \frac{1}{2\sigma_2^2} \sum_{t=t_0+1}^{n} (y_t - \boldsymbol{X}_t' \boldsymbol{\beta}_2)^2\right\} \tag{1-4}$$

t_0 の推定は，t_0 を $1, 2, \cdots n$ にわたって変化させ，尤度 $L(y/t_0)$ を最大とする時点 t_0^* を選定することによって行う。また，構造変化が発生しないという帰無仮説の検定のための統計量として，尤度比

$$\mu = \hat{\sigma}_1^{t_0} \hat{\sigma}_2^{(n-t_0)} / \hat{\sigma}^n \tag{1-5}$$

を用いる。ここで，$\hat{\sigma}$ は全標本期間にわたる回帰式の残差の推定値，$\hat{\sigma}_1$，$\hat{\sigma}_2$ は推定した t_0^* 前後のそれぞれの回帰式の残差の推定値である。

Brown, Durbin and Evance〔1〕は "recursive residuals" を用いて，構造変化時点の推定と検定についてより精緻な方法を提示している。いま，$\boldsymbol{b}_t$ を回帰係数ベクトル $\boldsymbol{\beta}$ の最初の K 標本（K は独立変数の数に一致させる）による推定値とし，X_t を行ベクトル $\boldsymbol{X}_1'$, $\boldsymbol{X}_2'$, $\cdots$, $\boldsymbol{X}_t'$ からなる独立変数行列とすると，"recursive residuals"，W_t は

$$W_t = \frac{y_t - \boldsymbol{X}_t' \boldsymbol{b}_{t-1}}{[1 + \boldsymbol{X}_t' (X_{t-1}' X_{t-1})^{-1} \boldsymbol{X}_t]^{1/2}}, \quad t = K+1, \cdots, n \tag{1-6}$$

と定義される。また，回帰係数ベクトルが標本期間にわたって一定，H_0：$\boldsymbol{\beta}_1 = \boldsymbol{\beta}_2 = \cdots = \boldsymbol{\beta}_n = \boldsymbol{\beta}$，という帰無仮説の下では，$W_t$ の系列は

$$\begin{aligned} &W_t \sim N(0, \sigma^2), \\ &E(W_t W_{t+1}) = 0, \quad t = K+1, \cdots, n \end{aligned} \tag{1-7}$$

となる。

$\boldsymbol{b}_t$ が t_0 時点まで一定であり，その後に変動が生じたとすると，W_t 系列は t_0 時点までは零レベルを中心に推移し，その後は零レベルから離れていく。W_t 系列が零レベルから離れた値で推移するときは $\boldsymbol{b}_t, t=K+1, \cdots, n$ が一定でないことを示し，特に W_t 系列の平均値の符合が変化した場合，係数ベクトルの不安定なパターンに変動が生じたことを示す。

帰無仮説 H_0 の統計的有意性の検定のための統計量として，Brown *et al.*〔1〕は W_t の累積和あるいは二乗累積和を提示しているが，これらの検出力が低いという指摘がある（たとえば，Garbade〔6〕参照)。

回帰係数ベクトルが標本期間において一定でない場合，W_t 系列間に相関が生じ，反対に一定のときには系列相関が零である。したがって，帰無仮説 H_0 の統計的検定は W_r 系列の相関をチェックすることによって行うことができる。その系列相関のチェックには W_t の vonNeumann 比を用いるのも一つの方法である①。

現在，構造変化の時点が不明な場合のスイッチング回帰モデルは，二つの回帰式の切換えは確率現象であるとする考え方にしたがって定式化されている。構造変化によって標本期間が二分され，各々の区間から観測される従属変数は独立変数によって決定されるとして二つの回帰式があてはめられる。最初の区間にあてはめた回帰式が出現する確率は λ，二番目の区間の回帰式が出現する確率は $(1-\lambda)$ である。しかし λ は未知であるが一定の確率と考えられている。

以下では，構造変化時点が不明な場合のスイッチング回帰モデルの定式化と，その推定の問題について考察する。スイッチング回帰モデルの分析には，Bayes 統計学による接近も試みられているが②，本稿での接近方法は Bayes 統計学のものではない。

注

① recursive residuals と vonNeumann 比を用いた構造変化時点の推定と統計的検定の問題に関しては，Harada, K. and R. Shirogane: "Recursive Estimation Analysis for Examining the Stability of Regression Relationships," *Memoirs of the Kokushikan University Computer Center*, No. 6 (1985), 1–20. を参照。

② Bayes 統計学の方法によるスイッチング回帰モデルの分析については，Holbert〔9〕を参照。

2 スイッチング回帰モデル

構造変化が発生しているということ以外に先験的情報がない場合，スイッチング回帰モデルは次のような考え方にしたがって定式化される。標本期間（$i=1,2,\cdots,n$）が構造変化時点（未知）によって二分され，未知のしかし一定の確率 λ で区間Iに於てあるいは確率 $(1-\lambda)$ で区間IIに於て，観測値（標本）が従属変数 y_i と独立変数ベクトル $\boldsymbol{X}_i$ について与えられる。この場合の回帰式は

$$
\begin{aligned}
y_i&=\boldsymbol{X}_i'\boldsymbol{\beta}_1+u_{1i}, \quad \text{確率 } \lambda,\\
y_i&=\boldsymbol{X}_i'\boldsymbol{\beta}_2+u_{2i}, \quad \text{確率 } (1-\lambda),
\end{aligned}
\tag{2-1}
$$

$\boldsymbol{\beta}_j=(\beta_{j1}\cdots\beta_{jK})'$，$j=1,2$ である。ここで，誤差項が相互に独立に，$u_{1i}\sim N(0,\sigma_1^2)$，$u_{2i}\sim N(0,\sigma_2^2)$ であると仮定すれば，y_i についての確率密度関数は Day〔3〕が示すように混合密度関数となって

$$
\begin{aligned}
f_i=f(y_i|\lambda,\boldsymbol{\beta}_1',\boldsymbol{\beta}_2',\sigma_1^2,\sigma_2^2)&=\frac{\lambda}{\sqrt{2\pi}\sigma_1}\exp\left\{-\frac{1}{2\sigma_1^2}(y_i-\boldsymbol{X}_i'\boldsymbol{\beta}_1)^2\right\}\\
&+\frac{(1-\lambda)}{\sqrt{2\pi}\sigma_2}\exp\left\{-\frac{1}{2\sigma_2^2}(y_i-\boldsymbol{X}_i'\boldsymbol{\beta}_2)^2\right\}
\end{aligned}
\tag{2-2}
$$

である。また，未知のパラメータについての尤度関数は

$$
\begin{aligned}
L=\prod_{i=1}^{n}f_i=\prod_{i=1}^{n}\Bigg[&\frac{\lambda}{\sqrt{2\pi}\sigma_1}\exp\left\{-\frac{1}{2\sigma_1^2}(y_i-\boldsymbol{X}_i'\boldsymbol{\beta}_1)^2\right\}\\
&+\frac{1-\lambda}{\sqrt{2\pi}\sigma_2}\exp\left\{-\frac{1}{2\sigma_2^2}(y_i-\boldsymbol{X}_i'\boldsymbol{\beta}_2)^2\right\}\Bigg]
\end{aligned}
\tag{2-3}
$$

である。

ところで，i のある値において，$y_i - X_i'\boldsymbol{\beta}_1 \equiv 0$，となるように $\boldsymbol{\beta}_1$ を常に選ぶことができる。この場合，$\sigma_1 \to 0$ としたときの尤度関数 L の動きを調べてみる。いま，$\sigma_1 \to 0$ のとき，(2-3) 式の i 番目の項は無限大に接近し（$y_i \fallingdotseq X_i'\boldsymbol{\beta}_1$ のときには (2-3) 式の和の部分の第一項は零），また同式の和の部分の第二項が存在するから i 番目でない他のすべての項は零でないことになって，Lは有界とならない。このように，パラメータ空間のある端点において，Lが有界でないことから，スイッチング回帰モデルのパラメータの推定に最尤法を採用することが適切でないと考えられ，次節に記述するような積率母関数法が開発された。

しかし，Kiefer〔10〕はスイッチング回帰モデルのパラメータの最尤推定量が有効推定量③ であることを証明している。最尤法を用いた場合に問題となるのは，非線型の最大化問題を解かなければならないことである。それは，対数尤度関数が

$$\ln L = \sum_{i=1}^{n} \ln\left[\frac{\lambda}{\sqrt{2\pi}\sigma_1}\exp\left\{-\frac{1}{2\sigma_1^2}\left(y_i - \sum_{j=1}^{K}\beta_{1j}x_{ji}\right)^2\right\} + \frac{1-\lambda}{\sqrt{2\pi}\sigma_2}\exp\left\{-\frac{1}{2\sigma_2^2}\left(y_i - \sum_{j=1}^{K}\beta_{2j}x_{ji}\right)^2\right\}\right] \tag{2-4}$$

となるから，未知のパラメータについての第一次導関数が非線型で与えられることによるものである。この場合には，Newton 法などの反復推定の方法を用いなければならない。

注

③ 推定量の分散が常に Cramér-Rao の不等式の下限，すなわち Fisher の情報行列の逆数，に一致するとき，その推定量を有効推定量 (efficient estimator) という。

3　積率母関数推定と最尤推定

（1）　積率母関数による推定

ある所定の値 t とランダム変数 y についての積率母関数，$E(e^{ty})$，は t およびパラメータの関数となる。大数の強法則により，e^{ty} の標本平均，$\frac{1}{n}\sum_{i=1}^{n} e^{t_j y_i}$，は $E(e^{t_j y_i})$ に確率1で収束するから，積率母関数は e^{ty} の標本平均によって推定できる。

Quandt and Ramsey〔16〕は，スイッチング回帰モデルの推定にあたって，Q 個の異なる値の t を選定し（Q はパラメータの数より大きくとる），Q 個の積率母関数の値とその理論値の差の二乗和を最小にするようなパラメータを求めることを示唆している。これが積率母関数推定量（以下では MGF 推定量）である。

スイッチング回帰モデル（2-1）から

$$
\begin{aligned}
&y_i \sim N(\boldsymbol{X}_i'\boldsymbol{\beta}_1, \sigma_1^2), \quad \text{確率 } \lambda \\
&y_i \sim N(\boldsymbol{X}_i'\boldsymbol{\beta}_2, \sigma_2^2), \quad \text{確率 } (1-\lambda)
\end{aligned}
\qquad (3\text{-}1)
$$

となり，未知のパラメータ・ベクトル

$$\theta = (\boldsymbol{\beta}_1', \boldsymbol{\beta}_2', \sigma_1^2, \sigma_2^2, \lambda)'$$

のディメンジョンは

$$M = 2K + 3$$

である。

いま，$Q \geqq M$ とし，Q 個の値 $t_1, \cdots, t_Q$ を定めると積率母関数は次のように定義される。

$$G(\theta, t_j, \boldsymbol{X}_i) = \lambda \exp(t_j \boldsymbol{X}_i' \boldsymbol{\beta}_1 + t_j^2 \sigma_1^2/2) + (1-\lambda)\exp(t_j \boldsymbol{X}_i' \boldsymbol{\beta}_2 + t_j^2 \sigma_2^2/2) \tag{3-2}$$

更に，次の諸量を定義する。

$$z_{ij} = \exp(t_j y_i), \tag{3-3}$$

$$\gamma_{ij} = z_{ij} - G(\theta, t_j, \boldsymbol{X}_i), \tag{3-4}$$

$$\bar{z}_j = \frac{1}{n} \sum_{i=1}^{n} z_{ij}, \tag{3-5}$$

$$\bar{\gamma}_j = \frac{1}{n} \sum_{i=1}^{n} \gamma_{ij} = \bar{z}_j - G(\theta, t_j, \boldsymbol{X}_i), \tag{3-6}$$

$$\bar{\gamma} = (\bar{\gamma}_1, \cdots, \bar{\gamma}_Q)', \tag{3-7}$$

$$(\Omega_i)_{jk} = G(\theta, t_j + t_k, \boldsymbol{X}_i) - G(\theta, t_j, \boldsymbol{X}_i) G(\theta, t_k, \boldsymbol{X}_i) \tag{3-8}$$

$$\bar{\Omega} = \frac{1}{n} \sum_{i=1}^{n} \Omega_i, \tag{3-9}$$

$$(i=1, \cdots, n;\ j, k=1, \cdots, Q).$$

Quandt and Ramsey〔16〕は次の二乗和

$$S_1 = \sum_{j=1}^{Q} \bar{\gamma}_j^2 \tag{3-10}$$

を最小化することにより，スイッチング回帰モデルのパラメータ・ベクトルの MGF 推定量（以下では MGF θ_1 と表示）を求め，また Kiefer〔10〕は次の二乗和

$$S_2 = \frac{1}{n} \sum_{j=1}^{Q} \sum_{i=1}^{n} \gamma_{ij}^2 \tag{3-11}$$

を最小化して MGF 推定量（以下では MGF θ_2）を得ている。S_1 を最小化した場合と S_2 を最小化した場合では異なった推定量を導出することになるが，いずれの推定量も一致推定量であることが証明されている。

Schmidt〔17〕はこれらの推定量の漸近分布を次のように与えている。まず MGF θ_1 については，

$$\sqrt{n}\,(\hat{\theta}_1-\theta)\rightarrow N(0,\Phi_1), \tag{3-12}$$

$$\Phi_1=\lim_{n\to\infty} n\left[\left(\sum_{t=1}^{n} A_t\right)'\left(\sum_{t=1}^{n} A_t\right)\right]^{-1}\left(\sum_{t=1}^{n} A_t\right)'\left(\sum_{t=1}^{n} \Omega_t\right)\left(\sum_{t=1}^{n} A_t\right)$$
$$\times\left[\left(\sum_{t=1}^{n} A_t\right)'\left(\sum_{t=1}^{n} A_t\right)\right]^{-1}, \tag{3-13}$$

であり，また MGF θ_2 については,

$$\sqrt{n}\,(\hat{\theta}_2-\theta)\rightarrow N(0,\Phi_2), \tag{3-14}$$

$$\Phi_2=\lim_{n\to\infty} n\left(\sum_{t=1}^{n} A_t'A_t\right)^{-1}\left(\sum_{t=1}^{n} A_t'\Omega_t A_t\right)\left(\sum_{t=1}^{n} A_t'A_t\right)^{-1}, \tag{3-15}$$

である。ここで，A_t は $Q\times M$ 行列で次のように定義される。

$$(A_t)_{j,k}=\frac{\partial G(\theta,t_j,\boldsymbol{X}_t)}{\partial\theta_k},\quad (j=1,\cdots,Q_t;\ k=1,\cdots,M), \tag{3-15}$$

しかし，Schmidt〔17〕はこれらの結果からでは，MGF θ_1 と MGF θ_2 の漸近的有効性の優劣の判定をつけられないとして，MGF 推定量の改良を試みている。MGF θ_2 を改良するために,

$$S_3=\bar{\gamma}'\Omega^{-1}\bar{\gamma} \tag{3-16}$$

を最小化して MGF θ_3 を得て，その漸近分布を

$$\sqrt{n}\,(\hat{\theta}_3-\theta)\rightarrow N(0,\Phi_3), \tag{3-17}$$

$$\Phi_3=\lim_{n\to\infty} n\left[\left(\sum_{t=1}^{n} A_t\right)'\left(\sum_{t=1}^{n} \Omega_t\right)^{-1}\left(\sum_{t=1}^{n} A_t\right)\right]^{-1}, \tag{3-18}$$

のように与えている。

ここで，行列（$\Sigma_t A_t$）と行列（$\Sigma_t\Omega_t$）についてみると，Φ_3 は最良線型不偏推定量の形式をもつ共分散行列であり，Φ_1 は最小二乗推定量の形式を有している。したがって，MGF θ_3 のほうが MGF θ_1 に比べて漸近的有効性の高い推定量となる。しかし，MGF θ_3 と MGF θ_2 の比較において，その優劣は明確でないとしている。

また，MGF θ_2 を改良するために，

$$\gamma_i=(\gamma_{i1},\gamma_{i2},\cdots,\gamma_{iQ})',\quad (i=1,2,\cdots,n) \tag{3-19}$$

を定義し，

$$S_4=\frac{1}{n}\sum_{i=1}^{n}\gamma'_{ii}\hat{\Omega}^{-1}\gamma_i \tag{3-20}$$

を最小化して MGF θ_4 を得ている。ただし，$\hat{\Omega}_i^{-1}$ は θ の一致推定量（これは MGF θ_1，MGF θ_2 あるいは MGF θ_3 による $\hat{\theta}$）を基に計算される。MGF θ_4 の漸近分布は

$$\sqrt{n}(\hat{\theta}_4-\theta)\to N(0,\Phi_4) \tag{3-21}$$

$$\Phi_4=\lim_{n\to\infty} n\left(\sum_{i=1}^{n}A_i'\Omega_i^{-1}A_i\right)^{-1} \tag{3-22}$$

のように与えられている。

以上から，Φ_4 を Φ_1,Φ_2 および Φ_3 と比較すると，MGF θ_4 が MGF θ_1，MGF θ_2 そして MGF θ_3 に比して漸近的有効な推定量であるとしている。

MGF 推定量は一致推定量であり， Schmidt によって改良されたものは漸近的有効性が向上されているが， t の値の選定についての客観的な基準がなく任意性が伴う。この点が MGF 推定の難点である。

（2） 最尤法による推定

スイッチング回帰モデル (2-1) のパラメータ・ベクトルを，$\theta=(\lambda,\beta_1',\beta_2',\sigma_1^2,\sigma_2^2)$ としまたパラメータ空間が次のように与えられているとものとする。

$$0\leq\lambda\leq 1,\quad -\infty<\beta_{1j}<\infty,\quad -\infty<\beta_{2j}<\infty,\quad j=1,\cdots,K$$
$$0<\sigma_1<\infty,\quad 0<\sigma_2<\infty$$

パラメータの真の値 θ_0 は，$\lambda=0$，$\lambda=1$，$\sigma_1=0$ あるいは $\sigma_2=0$ を含まない閉領域 $\bar{R}$ の中に含まれていると仮定する。

（2-1）の従属変数の確率密度関数は（2-2）により，

$$f_t=f(y_t/\theta)=\frac{\lambda}{\sqrt{2\pi}\sigma_1}\exp\left\{-\frac{1}{2\sigma_1{}^2}(y_t-\boldsymbol{X}_t'\boldsymbol{\beta}_1)^2\right\}$$
$$+\frac{(1-\lambda)}{\sqrt{2\pi}\sigma_2}\exp\left\{-\frac{1}{2\sigma_2{}^2}(y_t-\boldsymbol{X}_t'\boldsymbol{\beta}_2)^2\right\}, \tag{3-23}$$

であるから，尤度方程式は

$$\frac{\partial \ln L}{\partial\theta}=0,\quad \ln L=\sum_{t=1}^{n}\ln f(y_t/\theta),$$

によって与えられる。

尤度関数（確率密度関数）$f(y_t/\theta)$ が Cramér の条件を満たしているならば，尤度方程式の唯一の解 θ_n（最尤推定量）が一致推定量であり，さらに，$\sqrt{n}(\theta_n-\theta_0)$ の分布は漸近的に正規分布となって，その平均が零，分散行列が $I(\theta_0)^{-1}$ である。なお，$I(\theta_0)$ は Fisher 情報行列を表わす④。

Kiefer〔10〕は $f(y_t/\theta)$ が Cramér の条件を満たすことを証明して，スイッチング回帰モデル（2-1）の最尤推定量が有効推定量であることを示している。

スイッチング回帰モデル（2-1）の尤度方程式は非線型となるから，パラメータの最尤推定量の導出は非線型の最適化問題を解くことになる。その解法の一つに Newton 法があり，この方法による反復計算では初期値（一致推定値）が必要となる。Kiefer〔10〕〔11〕はその初期値として，MGF 法による推定値（これは一致推定値である）を用いることを奨めている。

Newton 法による尤度方程式の解（極大値）を求めるための反復計算ステップは

$$\theta_{h+1}=\theta_h-H_h^{-1}F_h,\quad h=0,1,2,\cdots, \tag{3-24}$$

$$F_h = \frac{\partial \ln L}{\partial \theta_h} = \sum_{t=1}^{n} \frac{\partial f_t}{\partial \theta_h} \cdot \frac{1}{f_t}, \tag{3-25}$$

$$H_h = \frac{\partial^2 \ln L}{\partial \theta_h \partial \theta_h'} = \sum_{t=1}^{n} \frac{\partial^2 f_t}{\partial \theta_h \partial \theta_h'} \cdot \frac{1}{f_t} - \sum_{t=1}^{n} \left(\frac{\partial f_t}{\partial \theta_h}\right)\left(\frac{\partial f_t}{\partial \theta_h'}\right)\frac{1}{f_t^2} \tag{3-26}$$

のように定義される。

Newton 法のプロセスでは，収束が得られるまで反復計算が行われる。パラメータ・ベクトル θ の系列は，尤度 L が極大となる値に収束する。この方法で求めた推定値は t の値に依存するものではなくなり，しかも漸近的有効性を有す⑤。対数尤度関数の第二次導関数の行列は，t の値に依存しない漸近的分散を与える。

注

④ Cramér〔2〕は単一パラメータの場合について最尤推定量の漸近特性を検討している。複数パラメータの場合の Cramér 条件を付録に示す。

最尤推定量の漸近特性については，Dhrymes〔4〕, chapter 3, あるいは Theil〔19〕, chapter 8 を参照。

⑤ 最尤推定量のもつ有効推定量としての特性が保たれている。

4 むすび

構造変化が起こる時点が明確でない場合のスイッチング回帰モデルのパラメータの推定については，最尤法による方式が Goldfeld and Quandt〔7〕, Hartley〔8〕および Kiefer〔10〕によって提示され，積率母関数を基にした推定方式が Quandt and Ramsey〔16〕および Schmidt〔17〕によって開発されている。

確率変数としての推定量の優劣を比較するために，推定量の分布の分布特性値を基準にしてその検討が行われる。この方法によると，両者ともにそれぞれの漸近的分布は正規分布となり，いずれも平均は零であるが，分

散行列に差異がある。最尤推定量の分散行列だけが Cramér-Rao の不等式の下限に一致し，最尤推定量が有効推定量となる。

積率母関数に基づく推定量は一致推定量となるが，有効推定量でなく，しかも積率母関数の t 値の選定に恣意性を伴う。この t 値の選定に客観的基準がなく，その値いかんによって推定量が影響を受ける。

スイッチング回帰モデル推定に最尤法を適用すると，その尤度方程式は非線型となるので，この非線型の最適化問題を解くには，Newton 法などによる反復計算が必要となる。この場合の初期値として，一致推定値である積率母関数による推定値を採用することが良策である。Newton 法の反復計算は尤度が極大となるまで行われるので，こうして求めた推定値は積率母関数の t 値に依存しないものとなり，また，最尤推定法のもつ漸近的有効性も保持される。

このように推定されたスイッチング回帰モデルは，確率 λ で標本期間内の領域 I に於て推定された構造（経済関係）であること，あるいは確率 $(1-\lambda)$ で領域 II に於て推定された構造が示されることになる。

ところで，本稿でのモデルは構造変化に関する先験的情報を定式化していないものであったが，構造変化についての先験的情報を，スイッチング回帰モデルの定式化に取り入れる試みがなされている。Lee and Porter〔13〕は，価格競争がある場合のカルテル行為の変化の分析に，スイッチング回帰モデルを用いている。そのモデルの定式化に構造変化についての先験的情報を取り入れている。価格競争が行われているとき，カルテルを結んでいる各企業が，協調的行動から非協調的行動に移行するという命題を，スイッチング回帰モデルとして表現するときに，構造変化時点に関しての先験的情報（この場合に不確定なものである）を遷移確率として定式化している。

構造変化時点（標本区分）に関しての不確定な先験的情報を取り入れた Lee and Porter のスイッチング回帰モデルについても，その推定においては，最尤法が採用されている。

付　最尤推定量の漸近的有効性の証明のための Cramér の条件

条件1：ランダム変数 y_i $(i=1,\cdots,n)$ および真のパラメーター・ベクトル θ_0 が含まれる空間 $\bar{R}$ に属するパラメーター・ベクトル θ について，密度関数 $L=\prod_{i=1}^{n}f(y_i,\theta)$ の導関数

$$\frac{\partial\ln L}{\partial\theta},\quad \frac{\partial^2\ln L}{\partial\theta\partial\theta'}\quad \text{および}\quad \frac{\partial^3\ln L}{\partial\theta\partial\theta'\partial\theta}$$

が存在する。

条件2：$y=(y_1\cdots y_n)$ および $\bar{R}$ に属する θ について，次の関係が成立する。

$$\left|\frac{\partial\ln L}{\partial\theta}\right|<F(y),\quad \left|\frac{\partial^2\ln L}{\partial\theta\partial\theta'}\right|<G(y)$$

$$\left|\frac{\partial^3\ln L}{\partial\theta\partial\theta'\partial\theta}\right|<H(y),\quad L=\sum_{i=1}^{n}f(y_i,\theta)$$

ここで，$H(y)$ は $\int_{-\infty}^{\infty}H(y)f(y,\theta)dy\leq M<\infty$ であり，また $F(y)$ と $G(y)$ は共に有界である。

条件3：$\bar{R}$ に属する θ について，次の行列（Fisher 情報行列）

$$I(\theta)=-E\left[\frac{\partial^2\ln L}{\partial\theta\partial\theta'}\right],\quad L=\sum_{i=1}^{n}f(y_i;\theta)$$

が正値定符号である。

参 考 文 献

〔1〕 Brown, R. L., J. Durbin and J. M. Evance: “Techniques for

Testing the Constancy of Regression Relationships over Time", *Journal of the Royal Statistical Society*, Ser. B 37 (1975), 149-163.

[2] Cramér, H.: *Mathematical Methods of Statistics.* Prinston: Princeton University Press, 1946.

[3] Day, N. E.: "Estimating the Components of a Mixture of Normal Distributions," *Biometrika*, 56 (1969), 463-474.

[4] Dhrymes, P. J.: *Econometrics: Statistical Foundations and Applications.* New York: Harper and Row, 1970.

[5] Fomby, T. B., R. C. Hill and S. R. Johnson: *Advanced Econometric Methods.* New York: Springer-Verlag, 1984.

[6] Galbade, K.: "Two Methods for Examining the Stability of Regression Coefficients," *Journal of the American Statistical Association*, 72 (1977), 54-63.

[7] Goldfeld, S. M. and R. E. Quandt: "Techniques for Estimating Switching Regressions," in *Studies in Nonlinear Estimation*, ed. by S. M. Goldfeld and R. E. Quandt. Cambridge, Massachusetts: Ballinger, 1976, 3-35.

[8] Hartley, M. J.: "Comment" (on [16]), *Journal of the American Statistical Association*, 73 (1978), 738-741.

[9] Holbert, D.: "A Bayesian Analysis of a Switching Linear Model," *Journal of Econometrics*, 19 (1982), 77-87.

[10] Kiefer, N. M.: "Discrete Parameter Variation: Efficient Estimation of a Switching Regression Model," *Econometrica*, 46 (1978), 427-434.

[11] ———: "Coment" (on [16]), *Journal of the American Statistical Association*, 73 (1978), 744-745.

[12] ———: "A Note on Switching Regressions and Logistic Discrimination," *Econometrica*, 48 (1980), 637-639.

[13] Lee, L. F. and R. H. Porter: "Switching Regression Models with Imperfect Sample Separation Information—with an Application on Cartel Stability," *Econometrica*, 52 (1984), 391-418.

[14] Poirier, D.: "Piecewise Regression Using Cubic Splines," *Journal of the American Statistical Association*, 68 (1973), 515-

524.

〔15〕 Quandt, R. E. : "A New Approach to Estimating Switching Regressions," *Journal of the American Statistical Association*, 67 (1972), 306-310.

〔16〕 Quandt, R. E. and J. Ramsey : "Estimating Mixtures of Normal Distributions and Switching Regressions," *Journal of the American Statistical Association*, 73 (1978), 730-738.

〔17〕 Schmidt, P. : "An Improved Version of the Quandt-Ramsey MGF Estimator for Mixtures of Normal Distributions and Switching Regressions," *Econometrica*, 50 (1982), 501-516.

〔18〕 Swamy, P. A. V. B. and J. S. Mehta : "Bayesian and Non-Bayesian Analysis of Switching Regressions and of Random Coefficient Regression Models," *Journal of the American Statistical Association*, 70 (1975), 539-602.

〔19〕 Theil, H. : *Principles of Econometrics*. Amsterdam : North-Holland, 1979.

（1985年6月）